

REMARKS

Applicant is in receipt of the Office Action mailed October 7, 2004. Claims 7-10, 15-28, 39-41, 48-51, 56-64, and 75-77 remain pending in the case. Further consideration of the present case is earnestly requested in light of the following remarks.

Section 103 Rejections

Claims 7-10, 15-28, 39-41, 48-51, 56-64, and 75-77 were rejected under 35 U.S.C. 103 (a) as being unpatentable over U.S. Patent Number 5,826,249 issued to Richard Skeirik ("Skeirik") in view of U.S. Patent Number 6,427,141 issued to Stephen Barnhill ("Barnhill"). Applicant respectfully traverses the rejection.

Claim 7 recites:

7. (Original) A method for training a support vector machine, the method comprising:

- (1) constructing a list containing at least two training sets;
- (2) training the support vector machine using said at least two training sets in said list;

- (3) constructing a new training set and replacing an oldest training set in said list with said new training set; and

- (4) repeating (2) and (3) at least once;

wherein at least one of (1) and (3) comprises:

- (a) retrieving training input data from a historical database, wherein said training input data has one or more timestamps;

- (b) selecting a training input data time period based on said one or more timestamps; and

- (c) retrieving an input data indicated by said training input data time period.

As the Examiner is certainly aware, to establish a prima facie obviousness of a claimed invention, all claim limitations must be taught or suggested by the prior art. In re

Royka, 490 F.2d 981, 180 U.S.P.Q. 580 (C.C.P.A. 1974), MPEP 2143.03. Obviousness cannot be established by combining or modifying the teachings of the prior art to produce the claimed invention, absent some teaching or suggestion or incentive to do so. In re Bond, 910 F.2d 81, 834, 15 USPQ2d 1566, 1568 (Fed. Cir. 1990).

Moreover, as held by the U.S. Court of Appeals for the Federal Circuit in *Ecolchem Inc. v. Southern California Edison Co.*, an obviousness claim that lacks evidence of a suggestion or motivation for one of skill in the art to combine prior art references to produce the claimed invention is defective as hindsight analysis.

In addition, the showing of a suggestion, teaching, or motivation to combine prior teachings “must be clear and particular Broad conclusory statements regarding the teaching of multiple references, standing alone, are not ‘evidence’.” In re Dembiczak, 175 F.3d 994, 50 USPQ2d 1614 (Fed. Cir. 1999). The art must fairly teach or suggest to one to make the specific combination as claimed. That one achieves an improved result by making such a combination is no more than hindsight without an initial suggestion to make the combination.

Applicant respectfully submits neither Skeirik nor Barnhill provides a motivation to combine.

For example, nowhere does Skeirik mention, or even hint at, support vector machines. Rather, Skeirik is *specifically* directed to *neural networks*, their training, and operation. In other words, the entire point of Skeirik is the training and operation of *neural networks*, not support vector machines. Applicant notes that neither the cited art, nor any other art Applicant is aware of, indicates that techniques applicable to neural networks are necessarily also applicable to support vector machines.

Applicant notes that if a proposed modification would render the prior art feature unsatisfactory for its intended purpose, then there is no suggestion or motivation to make the proposed modification. In re Gordon, 733 F.2d 900 (Fed. Cir. 1984). Since Skeirik is particularly directed to the training of neural networks, replacing Skeirik’s neural network with a support vector machine (or a plurality of support vector machines as taught by Barnhill) would certainly render Skeirik’s technique unsatisfactory for training

a neural network. Thus, Applicant respectfully submits that Skeirik does not provide a motivation to combine.

Similarly, Barnhill is specifically directed to " A system and method for enhancing knowledge discovery from data using multiple learning machines in general and multiple support vector machines in particular." (Abstract). In other words, Barnhill discloses a system and method for training and using a plurality of learning machines, and is specifically directed to the training and operation of multiple support vector machines. Applicant notes that each of the multiple support vector machines has a different "kernel" which governs operation and utility of that machine, and that the user tests each machine against training data and picks the machine which provides the best solution. While Barnhill does indicate that in some embodiments the multiple machines may include neural networks, Applicant notes that Barnhill also provides many reasons for *not* using neural networks. For example, in col. 1 line 53 – col. 2, line 12, Barnhill states:

However, there are various problems with back-propagation neural network approaches that prevent neural networks from being well-controlled learning machines. For example, a significant drawback of back-propagation neural networks is that the empirical risk function may have many local minimums, a case that can easily obscure the optimal solution from discovery by this technique. Standard optimization procedures employed by back-propagation neural networks may converge to a minimum, but the neural network method cannot guarantee that even a localized minimum is attained much less the desired global minimum. The quality of the solution obtained from a neural network depends on many factors. In particular the skill of the practitioner implementing the neural network determines the ultimate benefit, but even factors as seemingly benign as the random selection of initial weights can lead to poor results. Furthermore, the convergence of the gradient based method used in neural network learning is inherently slow. A further drawback is that the sigmoid function has a scaling factor, which affects the quality of approximation. Possibly the largest limiting

factor of neural networks as related to knowledge discovery is the "curse of dimensionality" associated with the disproportionate growth in required computational time and power for each additional feature or dimension in the training data.

The shortcomings of neural networks are overcome using support vector machines.

Thus, Barnhill clearly indicates that the use of multiple support vector machines in his disclosed system and method is preferred, and actively discourages use of neural networks, although allowing for the possibility. In other words, per Barnhill, support vector machines and neural networks are generally not interchangeable technologies. Applicant submits that due to Barnhill's enumeration of problems with neural networks, one of ordinary skill in the art would not be compelled to consider techniques generally applicable to neural networks to necessarily apply to support vector machines, and thus would not be motivated by Barnhill to combine Barnhill and Skeirik.

Moreover, Applicant notes that the only mention Barnhill makes of historical data is with respect to example sources of training or test data. Nowhere does Barnhill describe or even hint at using a list or sequence of multiple historical data sets where the list or sequence is updated by dropping the oldest data set and adding a new (younger) data set. Thus, Applicant respectfully submits that Barnhill does not provide a motivation to combine.

Applicant notes that the only motivation suggested by the Examiner to combine Skeirik and Barnhill is "because support vector machines of Barnhill provide a desirable solution for the problem of discovering knowledge from vast amount of input data" (*sic*). Applicant submits that this well-known utility of support vector machines in no way motivates the specific combination of Skeirik's historical data based neural network training technique with Barnhill's (multiple) support vector machines.

Thus, for at least the reasons provided above, Applicant respectfully submits that the Examiner's attempted combination of Skeirik and Barnhill is improper.

Furthermore, Applicant submits that even were Skeirik and Barnhill properly combinable, which Applicant argues they are not, the resulting combination would still not produce Applicant's invention as claimed.

The Examiner attempts to combine Skeirik and Barnhill by selecting particular portions of each while omitting other critical features of each. For example, the Examiner has apparently selected Skeirik's historical data based neural network training technique but has omitted the Skeirik's neural network, and has selected a single one of Barnhill's multiple support vector machines while omitting the other support vector machines, as well as Barnhill's disclosed (and claimed) technique for training (including optimization), and using the plurality of support vector machines.

Applicant notes that, per *Interconnect Planning Corp. v. Feil*, 774 F.2d 1132, 1143, 227 USPQ 543, 551 (Fed. Cir. 1985), it is insufficient to select from the prior art the separate components of the inventor's combination, using the blueprint supplied by the inventor. Applicant submits that the Examiner has simply selected particular portions or components of Skeirik and Barnhill based on Applicant's claims in an attempt to reconstruct Applicant's invention as claimed, which is clearly hindsight analysis.

Applicant submits that neither Skeirik nor Barnhill teaches or suggests all the limitations of claim 7, and so for at least the reasons provided above, claim 7 and those claims dependent therefrom are patentably distinct and non-obvious over Skeirik and Barnhill, taken singly or in combination, and are thus allowable.

Claim 18 recites:

18. (Original) A computer support vector machine process control method adapted for predicting output data provided to a controller used to control a process for producing a product having at least one product property, the computer support vector machine process control method comprising:

a processor;

a memory medium coupled to the processor, wherein the memory medium stores a support vector machine software program, wherein the support vector machine software program comprises:

- (1) monitoring for the availability of new training input data by monitoring for a change in an associated timestamp of said training input data;
- (2) constructing a training set by retrieving first input data corresponding to said training input data;
- (3) training the support vector machine using said training set; and
- (4) predicting the output data from second input data using the support vector machine.

Applicant submits that Skeirik and Barnhill have been shown above to *not* be properly combinable, and so further submits that these references may not be combined in an attempt to produce the invention as represented in claim 18. Moreover, Applicant notes that Barnhill nowhere teaches or suggests the limitations and features of claim 18. For example, Barnhill fails to disclose or suggest monitoring for new data based on timestamps associated with the data, and using the new data to train a support vector machine. Nor does Skeirik teacher suggest, or even hint at, using the features and limitations of claim 18 to train and operate a support vector machine. Thus, for similar reasons as above, Applicant submits that Skeirik and Barnhill, taken singly or in combination, fail to teach or suggests all the features and limitations of claim 18. Thus claim 18 and those claims dependent therefrom are patentably distinct and nonobvious over Skeirik and Barnhill, and are thus allowable.

Independent claim 24 includes similar limitations as claim 18, and so the above arguments apply with equal force to this claim. Thus, for at least the reasons provided above, claim 24 and those claims dependant therefrom are patently distinct and non-obvious over Skeirik and Barnhill and are thus allowable.

Claim 15 recites:

15. (Original) A method for constructing training sets for a support vector machine, the method comprising:

- (1) developing a first training set for a support vector machine by:
 - (a) retrieving first training input data from a historical database, wherein said first training input data has a first one or more timestamps;

(b) selecting a first training input data time period based on said first one or more timestamps; and

(c) retrieving first input data indicated by said first training input data time period; and

(2) developing a second training set for said support vector machine by:

(a) retrieving second training input data from said historical database, wherein said second training input data has a second one or more timestamps;

(b) selecting a second training input data time period based on said second one or more timestamps; and

(c) retrieving second input data indicated by said second training input data time period.

Claim 23 recites:

23. (Original) A method for constructing training sets for a support vector machine, the method comprising:

(a) retrieving training input data from a historical database, wherein said training input data has one or more timestamps;

(b) selecting a training input data time period based on said one or more timestamps; and

(c) retrieving input data indicated by said training input data time period.

Claim 39 recites:

39. (Original) A method for training a support vector machine used to control a process, the method comprising:

building a first training set using training data, wherein said training data includes one or more timestamps indicating a chronology of said training data and one or more process parameter values corresponding to each timestamp, and wherein said first training set comprises process parameter values corresponding to a first time period in said chronology;

training a support vector machine using said first training set.

As may be seen, each of claims 15, 23, and 39, involves the use of timestamped data in training a support vector machine. In other words, chronologically ordered data are used to train a support vector machine. Nowhere does Barnhill teach or suggest this feature. Similarly, as argued above, Skeirik fails to teach or suggest, or even hint at, using the disclosed neural network training techniques (using timestamp data) to train a support vector machine. Additionally, as argued above, Skeirik and Barnhill are not properly combinable, and so further submits that these references may not be combined in an attempt to produce the invention as represented in any of claims 15, 23, and 39.

Thus, for at least the reasons presented above, claims 15, 23, and 39, and those claims respectively dependent therefrom, are patently distinct and non-obvious over Skeirik and Barnhill, taken singly or in combination, and are thus allowable.

Removal of the 103 rejection of claims 7-10, 15-28, 39-41, 48-51, 56-64, and 75-77 is respectfully requested.

Applicant also asserts that numerous ones of the dependent claims recite further distinctions over the cited art. However, since the independent claims have been shown to be patentably distinct, a further discussion of the dependent claims is not necessary at this time.

CONCLUSION

Applicant submits the application is in condition for allowance, and an early notice to that effect is requested.

If any extensions of time (under 37 C.F.R. § 1.136) are necessary to prevent the above referenced application(s) from becoming abandoned, Applicant(s) hereby petition for such extensions. If any fees are due, the Commissioner is authorized to charge said fees to Meyertons, Hood, Kivlin, Kowert & Goetzel PC Deposit Account No. 50-1505/5650-02200/JCH.

Also enclosed herewith are the following items:

☒ Return Receipt Postcard

Respectfully submitted,



Jeffrey C. Hood
Reg. No. 35,198
ATTORNEY FOR APPLICANT(S)

Meyertons, Hood, Kivlin, Kowert & Goetzel PC
P.O. Box 398
Austin, TX 78767-0398
Phone: (512) 853-8800
Date: 1/7/2005 JCH/MSW